



Qualitative Concurrent Stochastic Games with Imperfect Information

Vincent Gripon, Olivier Serre

► To cite this version:

Vincent Gripon, Olivier Serre. Qualitative Concurrent Stochastic Games with Imperfect Information. Automata, Languages and Programming, 36th International Colloquium, ICALP 2009, Jul 2009, Rhodes, Greece. pp.200-211. hal-00360857v3

HAL Id: hal-00360857

<https://hal.science/hal-00360857v3>

Submitted on 30 Aug 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Qualitative Concurrent Stochastic Games with Imperfect Information^{*}

Vincent Gripon and Olivier Serre

LIAFA (CNRS & Université Paris Diderot – Paris 7)

Abstract. We study a model of games that combines concurrency, imperfect information and stochastic aspects. Those are finite states games in which, at each round, the two players choose, *simultaneously* and *independently*, an action. Then a successor state is chosen accordingly to some fixed probability distribution depending on the previous state and on the pair of actions chosen by the players. Imperfect information is modeled as follows: both players have an equivalence relation over states and, instead of observing the exact state, they only know to which equivalence class it belongs. Therefore, if two partial plays are indistinguishable by some player, he should behave the same in both of them. We consider reachability (does the play eventually visit a final state?) and Büchi objective (does the play visit infinitely often a final state?). Our main contribution is to prove that the following problem is complete for 2-EXPTIME: decide whether the first player has a strategy that ensures her to almost-surely win against *any* possible strategy of her opponent. We also characterise those strategies needed by the first player to almost-surely win.

1 Introduction

Perfect information turn based two-player games on a graph [12] are widely studied in computer science. Indeed, they are a useful tool for both theoretical (for instance the modern proofs of Rabin’s complementation lemma rely on the memoryless determinacy of parity games [13]) and more practical applications. On the practical side, a major application of games is for the verification of reactive open systems. Those are systems composed of both a program and some (possibly hostile) environment. The verification problem consists of deciding whether the program can be restricted so that the system meets some given specification whatever the environment does. Here, restricting the system means synthesizing some controller, which, in term of games, is equivalent to designing a winning strategy for the player modeling the program [17].

The perfect information turn-based model, even if it suffices in many situations, is somewhat weak for the following two reasons. First, it does not permit to capture the behavior of real concurrent models where, in each step, the program and its environment *independently* choose moves, whose *parallel* execution

^{*} Supported by the ANR project JADE and by the ESF project GASICS.

determines the next state of the system. Second, in this model both players have, at each time, a perfect information on the current state of the play: this, for instance, forbids to model a system where the program and the environment share some public variables while having also their own private variables [18].

In this paper, we remove those two restrictions by considering concurrent stochastic games with imperfect information. Those are finite states games in which, at each round, the two players choose *simultaneously* and *independently* an action. Then a successor state is chosen accordingly to some fixed probability distribution depending on the previous state and on the pair of actions chosen by the players. Imperfect information is modeled as follows: both players have an equivalence relation over states and, instead of observing the exact state, they only see to which equivalence class it belongs. Therefore, if two partial plays are indistinguishable by some player, he should behave the same in both of them. Note that this model naturally captures several models studied in the literature [1, 11, 7, 8]. The winning conditions we consider here are reachability (is there a final state eventually visited?), Büchi (is there a final state that is visited infinitely often?) and their dual versions, safety and co-Büchi.

We study qualitative properties of those games (note that quantitative properties — *e.g.* deciding whether the value of the game is above a given threshold — are already undecidable in much weaker models [16]). More precisely, we investigate the question of deciding whether some player can almost-surely win, that is whether he has a strategy that wins with probability 1 against any counter strategy of the opponent. Our main contribution is to prove that, for both reachability and Büchi objectives, one can decide, in doubly exponential time (which is proved to be optimal), whether the first player has an almost-surely winning strategy. Moreover, when it is the case, we are also able to construct such a finite-memory strategy. We also provide intermediate new results concerning positive winning in safety (and co-Büchi) $1\frac{1}{2}$ -player games (*a.k.a* partial observation Markov decision process).

Related work. Concurrent games with perfect information have been deeply investigated in the last decade [2, 1, 7]. Games with imperfect information have been considered for turn-based model [18] as well as for concurrent models with only one imperfectly informed player [11, 8]. To our knowledge, the present paper provides the first positive results on a model of games that combines concurrency, imperfect information (on both sides) and stochastic transition function. In a recent independent work [4], Bertrand, Genest and Gimbert obtain similar results than the one presented here for a closely related model. Bertrand *et al.* also discuss qualitative determinacy results and consider the case where a player is more informed than the other. We refer the reader to [4] for a detailed exposition.

Comparison with our previous work [14]. Krishnendu Chatterjee and Laurent Doyen pointed out an important mistake in a previous version of this work originally published in ICALP'09 [14]. Indeed, in this previous work we additionally assumed that the players did not observe the actions they played and we proposed a variation of the constructions for this richer setting. It turned out

that the constructions were wrong. We refer to [9] for more insight on this point as well as for an answer to this more general problem.

2 Definitions

A **probability distribution** over a finite set X is a mapping $d : X \rightarrow [0, 1]$ such that $\sum_{x \in X} d(x) = 1$. In the sequel we denote by $\mathcal{D}(X)$ the set of probability distributions over X .

Given some set X and some equivalence relation \sim over X , $[x]_\sim$ stands for the equivalence class of x for \sim and $X/\sim = \{[x]_\sim \mid x \in X\}$ denotes the set of equivalence classes of \sim .

For some finite alphabet A , A^* (*resp.* A^ω) designates the set of finite (*resp.* infinite) words over A .

2.1 Arenas

A **concurrent arena with imperfect information** is a tuple $\mathcal{A} = \langle S, \Sigma_E, \Sigma_A, \delta, \sim_E, \sim_A \rangle$ where

- S is a finite set of control states;
- Σ_E (*resp.* Σ_A) is the (finite) set of actions for Eve (*resp.* Adam);
- $\delta : S \times \Sigma_E \times \Sigma_A \rightarrow \mathcal{D}(S)$ is the transition (total) function;
- \sim_E and \sim_A are two equivalence relations over states.

A play in a such an arena proceeds as follows. First it starts in some initial state s . Then Eve picks an action $\sigma_E \in \Sigma_E$ and, *simultaneously* and *independently*, Adam chooses an action $\sigma_A \in \Sigma_A$. Then a successor state is chosen accordingly to the probability distribution $\delta(s, \sigma_E, \sigma_A)$. Then the process restarts: the players choose a new pair of actions that induces, together with the current state, a new state and so on forever. Hence a **play** is an infinite sequence $s_0(\sigma_{E,0}, \sigma_{A,0})s_1(\sigma_{E,1}, \sigma_{A,1})s_2(\sigma_{E,2}, \sigma_{A,2}) \cdots$ in $(S \cdot (\Sigma_E \times \Sigma_A))^\omega$ such that for every $i \geq 0$, $\delta(s_i, \sigma_{E,i}, \sigma_{A,i})(s_{i+1}) > 0$. In the sequel we refer to a prefix of a play (ending in a state) as a **partial play** and we denote by $Plays(\mathcal{A})$ the set of all plays in arena \mathcal{A} .

The intuitive meaning of \sim_E (*resp.* \sim_A) is that two states s_1 and s_2 such that $s_1 \sim_E s_2$ (*resp.* $s_1 \sim_A s_2$) cannot be distinguished by Eve (*resp.* by Adam). We easily extend the relation \sim_E to partial plays (here Eve observes her actions, but does not observe Adam's actions): let $\lambda = s_0(\sigma_{E,0}, \sigma_{A,0})s_1(\sigma_{E,1}, \sigma_{A,1}) \cdots s_n$ and $\lambda' = s'_0(\sigma'_{E,0}, \sigma'_{A,0})s'_1(\sigma'_{E,1}, \sigma'_{A,1}) \cdots s'_n$ be two partial plays, then $\lambda \sim_E \lambda'$ if and only if $s_i \sim_E s'_i$ and $\sigma_{E,i} = \sigma'_{E,i}$ for all $i = 0, \dots, n$.

Note that perfect information concurrent arenas (in the sense of [2, 1]) correspond to the special case where \sim_E and \sim_A are the equality relation over S .

2.2 Strategies

In order to choose their moves the players follow strategies, and, for this, they may use all the information they have about what was played so far. However, if two partial plays are equivalent for \sim_E , then Eve cannot distinguish them, and should therefore behave the same. This leads to the following notion.

An **observation-based strategy** for Eve is a function $\varphi_E : (S/\sim_E) \cdot (\Sigma_E \cdot S/\sim_E)^* \rightarrow \mathcal{D}(\Sigma_E)$, *i.e.*, to choose her next action, Eve considers the sequence of observations she got so far. In particular, a strategy φ_E is such that $\varphi_E(\lambda) = \varphi_E(\lambda')$ whenever $\lambda \sim_E \lambda'$. Observation-based strategies for Adam are defined similarly.

Of special interest are those strategies that does not require memory: a **memoryless observation-based strategies** for Eve is a function from $S/\sim_E \rightarrow \mathcal{D}(\Sigma_E)$, that is to say these strategies only depend of the current equivalence class.

A **uniform strategy** for some player X is a strategy φ such that for all partial play λ , the probability measure $\varphi(\lambda)$ is uniform, *i.e.*, for all action $\sigma_X \in \Sigma_X$, either $\varphi(\lambda)(\sigma_X) = 0$ or $\varphi(\lambda)(\sigma_X) = \frac{1}{|\{\sigma_X \in \Sigma_X \mid \varphi(\lambda)(\sigma_X) \neq 0\}|}$. The set of memoryless uniform strategies for X is a finite set containing $(2^{|\Sigma_X|} - 1)^{|\Sigma|}$ elements. Equivalently those strategies can be seen as functions to (non-empty) sets of (authorised) actions.

A **finite-memory strategy** for Eve with memory M (M being a finite set) is some triple $\varphi = (Move, Up, m_0)$ where $m_0 \in M$ is the initial memory, $Move : M \rightarrow \mathcal{D}(\Sigma_E)$ associates a distribution of actions with any element in the memory M and $Up : M \times S/\sim_E \times \Sigma_E \rightarrow M$ is a mapping updating the memory with respect to some observation (and the last action played by Eve). One defines $\varphi(s_0) = Move(m_0)$ and $\varphi(s_0(\sigma_{E,0}, \sigma_{A,0}) \cdots s_n) = Move(Up(\cdots Up(Up(m_0, [s_1]/\sim_E, \sigma_{E,0}), [s_2]/\sim_E, \sigma_{E,1}), \cdots, [s_n]/\sim_E, \sigma_{E,n-1}) \cdots))$ for any $n \geq 1$. Hence, a finite-memory strategy is some observation-based strategy that can be implemented by a finite transducer whose set of control states is M .

2.3 Probability Space and Outcomes of Strategies

Let $\mathcal{A} = \langle S, \Sigma_E, \Sigma_A, \delta, \sim_E, \sim_A \rangle$ be a concurrent arena with imperfect information, let $s_0 \in S$ be an initial state, φ_E be a strategy for Eve and φ_A be a strategy for Adam. In the sequel we are interested in defining the probability of a (measurable) set of plays knowing that Eve (*resp.* Adam) plays accordingly φ_E (*resp.* φ_A). This is done in the classical way: first one defines the probability measure for basic sets of plays (called here *cones* and corresponding to plays having some initial common prefix) and then extends it in a unique way to all measurable sets.

First define $Outcomes(s_0, \varphi_E, \varphi_A)$ to be the set of all possible plays when the game starts on s_0 and when Eve and Adam plays respectively accordingly to φ_E and φ_A . More formally, an infinite play $\lambda = s_0(\sigma_{E,0}, \sigma_{A,0})s_1(\sigma_{E,1}, \sigma_{A,1})s_2 \cdots$ belongs to $Outcomes(s_0, \varphi_E, \varphi_A)$ if and only if, for every $i \geq 0$, $\varphi_E(s_0\sigma_{E,0}s_1\sigma_{E,1} \cdots s_i)(\sigma_{E,i}) >$

0 and $\varphi_A(s_0\sigma_{A,0}s_1\sigma_{A,1}\cdots s_i)(\sigma_{A,i}) > 0$ (i.e. σ_X is possible accordingly to φ_X , for $X = E, A$).

Now, for any partial play λ , the **cone** for λ is the set $\text{cone}(\lambda) = \lambda \cdot ((\Sigma_E \times \Sigma_A) \cdot S)^\omega$ of all infinite plays with prefix λ . Denote by Cones the set of all possible cones and let \mathcal{F} be the Borel σ -field generated by Cones considered as a set of basic open sets (i.e. \mathcal{F} is the smallest set containing Cones and closed under complementation, countable union and countable intersection). Then $(\text{Plays}(\mathcal{A}), \mathcal{F})$ is a σ -algebra.

A pair of strategies (φ_E, φ_A) induces a probability space over $(\text{Plays}(\mathcal{A}), \mathcal{F})$. Indeed one can define a measure $\mu_{s_0}^{\varphi_E, \varphi_A} : \text{Cones} \rightarrow [0, 1]$ on cones (this task is easy as a cone is uniquely defined by a finite partial play) and then uniquely extend it to a probability measure on \mathcal{F} using the Carathéodory Unique Extension Theorem. For this, one defines $\mu_{s_0}^{\varphi_E, \varphi_A}$ inductively on cones:

- $\mu_{s_0}^{\varphi_E, \varphi_A}(\text{cone}(s)) = 1$ if $s = s_0$ and $\mu_{s_0}^{\varphi_E, \varphi_A}(s) = 0$ otherwise.
- For every partial play λ ending in some vertex s ,

$$\mu_{s_0}^{\varphi_E, \varphi_A}(\text{cone}(\lambda \cdot (\sigma_E, \sigma_A) \cdot s')) = \mu_{s_0}^{\varphi_E, \varphi_A}(\text{cone}(\lambda)) \cdot \varphi_E(\lambda)(\sigma_E) \cdot \varphi_A(\lambda)(\sigma_A) \cdot \delta(s, \sigma_E, \sigma_A)(s')$$

Denote by $\text{Pr}_{s_0}^{\varphi_E, \varphi_A}$ the unique extension of $\mu_{s_0}^{\varphi_E, \varphi_A}$ to a probability measure on \mathcal{F} . Then $(\text{Plays}(\mathcal{A}), \mathcal{F}, \text{Pr}_{s_0}^{\varphi_E, \varphi_A})$ is a probability space.

2.4 Objectives, Value of a Game

Fix a concurrent arena with imperfect information \mathcal{A} . An objective for Eve is a measurable set $\mathcal{O} \subseteq \text{Plays}(\mathcal{A})$: a play is won by her if it belongs to \mathcal{O} ; otherwise it is won by Adam. A **concurrent game with imperfect information** is a triple $(\mathcal{A}, s_0, \mathcal{O})$ where \mathcal{A} is a concurrent arena with imperfect information, s_0 is an initial state and \mathcal{O} is an objective. In the sequel we focus on the following special classes of objectives (note that all of them are Borel sets hence measurable) that we define as means of a subset $F \subseteq S$ of **final states**.

- **Reachability objective**: a play is winning if it eventually goes through some final state.
- **Safety objective**: a play is winning if it never goes through a final state.
- **Büchi objective**: a play is winning if it goes infinitely often through final states.
- **Co-Büchi objective**: a play is winning if it goes finitely often through final states.

A reachability (resp. safety, Büchi, co-Büchi) game is a game equipped with a reachability (resp. safety, Büchi, co-Büchi) objective. In the sequel we may replace \mathcal{O} by F when it is clear from the context which winning condition we consider.

Fix a concurrent game with imperfect information $\mathbb{G} = (\mathcal{A}, s_0, \mathcal{O})$. A strategy φ_E for Eve is **almost-surely winning** if, for any counter-strategy φ_A for Adam, $\text{Pr}_{s_0}^{\varphi_E, \varphi_A}(\mathcal{O}) = 1$. If such a strategy exists, we say that Eve **almost-surely wins** \mathbb{G} . A strategy φ_E for Eve is **positively winning** if, for any counter-strategy φ_A for Adam, $\text{Pr}_{s_0}^{\varphi_E, \varphi_A}(\mathcal{O}) > 0$. If such a strategy exists, we say that Eve **positively wins** \mathbb{G} .

3 Knowledge Arena

For the rest of this section we let \mathcal{A} be a concurrent arena with imperfect information with $\mathcal{A} = \langle S, \Sigma_E, \Sigma_A, \delta, \sim_E, \sim_A \rangle$ and let $s_0 \in S$ be some initial state.

Even, if she does not see the precise control state, Eve can deduce information about it from previous information on the control state and from the action she just played. We should refer to this as the **knowledge** of Eve, which formally is a set of states. Assume Eve knows that the current state belongs to some set $K \subseteq S$. After the next move Eve observes the equivalence class $[s]_{\sim_E}$ of the new control state and she also knows that she played action σ_E : hence she can compute the set of possible states the play can be in. This is done using the function $\text{UpKnow} : 2^S \times [S]_{\sim_E} \times \Sigma_E \rightarrow 2^{\Sigma^S}$ defined by letting

$$\text{UpKnow}(K, [s]_{\sim_E}, \sigma_E) = \{t \sim_E s \mid \exists r \in K, \exists \sigma_A \in \Sigma_A \text{ s.t. } \delta(r, \sigma_E, \sigma_A)(t) > 0\}$$

i.e. in order to update her current knowledge, observing in which equivalence class is the new control state, and knowing that she played σ_E , Eve computes the set of all states in this class that may be reached from a state in her former knowledge.

Finally, we let

$$\begin{aligned} \text{Knowledge}(s_0 \sigma_{E,0} s_1 \cdots s_n) = \\ \text{UpKnow}(\text{UpKnow}(\cdots \text{UpKnow}(\{s_0\}, [s_1]_{\sim_E}, \sigma_{E,1}) \cdots), [s_n]_{\sim_E}, \sigma_{E,n}) \end{aligned}$$

Based on our initial remark and on the notion of knowledge we define the **knowledge arena associated with \mathcal{A}** , denoted \mathcal{A}^K . The arena \mathcal{A}^K is designed to make explicit the information Eve can collect about the possible current state (*i.e.* the knowledge). We define $\mathcal{A}^K = \langle S^K, \Sigma_E, \Sigma_A, \delta^K, \sim_E^K, \sim_A^K \rangle$ as follows:

- $S^K = \{(s, K) \in S \times 2^S \mid K \subseteq [s]_{\sim_E}\}$: the first component is the real state and the second one is the current knowledge of Eve;
- $\delta^K((s, K), \sigma_E, \sigma_A)(s', K') = 0$ if $K' \neq \text{UpKnow}(K, [s']_{\sim_E}, \sigma_E)$; and $\delta^K((s, K), \sigma_E, \sigma_A)(s', K') = \delta(s, \sigma_E, \sigma_A)(s')$ otherwise: δ^K behaves as δ on the first components and *deterministically* updates the knowledge;
- $(s, K) \sim_E^K (s', K'')$ if and only if $K = K'$ (implying $s \sim_E s'$): Eve only observes her knowledge;
- $(s, K) \sim_A^K (s', K')$ if and only if $s \sim_A s'$: Adam does not observe Eve's knowledge.

Consider an observation-based strategy φ for Eve in the arena \mathcal{A} . Then it can be converted into an observation-based strategy on the associated knowledge arena. For this, remark that in the knowledge arena, those states reachable from the initial state $(s_0, \{s_0\})$ are of the form (s, K) with all states in K being equivalent with s with respect to \sim_E . Then one can define

$$\varphi^K((s_0, K_0) \sigma_{E,0} (s_1, K_1) \sigma_{E,1} \cdots (s_n, K_n)) = \varphi([s_0]_{\sim_E} \sigma_{E,0} [s_1]_{\sim_E} \sigma_{E,1} \cdots [s_n]_{\sim_E})$$

Note that φ^K is observation-based as, for all $0 \leq h \leq n$, $[s_h]_{\sim_E}$ is uniquely defined from the K_h , that are observed by Eve in the knowledge arena.

Conversely, any observation-based strategy in the knowledge arena can be converted into an observation-based strategy in the original arena. Indeed, consider some observation-based strategy φ^K in the knowledge arena: it is a mapping from $2^S \cdot (\Sigma_E \cdot 2^S)^*$ into $\mathcal{D}(\Sigma_E)$ (the equivalent classes of the relation \sim_E^K are, by definition, isomorphic with 2^S). Now, note that Eve can, while playing in \mathcal{A} , compute on the fly her current knowledge (applying function UpKnow to her previous knowledge and to the last action she played): hence along a play $s_0(\sigma_{E,0}, \sigma_{A,0})s_1(\sigma_{E,1}, \sigma_{A,1}) \cdots s_n$ she can compute the corresponding sequence $K_0 K_1 \cdots K_n$ of knowledges. Now it suffices to consider the observation-based strategy φ for Eve in the initial arena defined by:

$$\varphi(s_0 \sigma_{E,0} s_1 \sigma_{E,1} \cdots s_n) = \varphi^K(K_0 \sigma_{E,0} K_1 \sigma_{E,1} \cdots K_n)$$

Note that this last transformation (taking a strategy φ^K and producing a strategy φ) is the inverse of the first transformation (taking a strategy φ and producing a strategy φ^K). In particular, it proves that the observation-based strategies in both arena are in bijection. It should be clear that those strategies for Adam in both games are the same (as what he observes is identical).

Assume that \mathcal{A} is equipped with a set F of final states. Then one defines the final states in \mathcal{A}^K by letting $F^K = \{(f, K) \mid f \in F\} \cap S^K$: this allows to define an objective \mathcal{O}^K in \mathcal{A}^K from an objective \mathcal{O} in \mathcal{A} . Based on the previous observations, we derive the following.

Proposition 1. *Let $\mathbb{G} = (\mathcal{A}, s_0, \mathcal{O})$ be some imperfect information game equipped with a reachability (resp. safety, Büchi, co-Büchi) objective. Let $\mathbb{G}^K = (\mathcal{A}^K, (s_0, \{s_0\}), \mathcal{O}^K)$ be the associated game played on the knowledge arena. Then for any strategies φ_E, φ_A for Eve and Adam, the following holds:*

$\text{Pr}_{s_0}^{\varphi_E, \varphi_A}(\mathcal{O}) = \text{Pr}_{(s_0, \{s_0\})}^{\varphi_E^K, \varphi_A^K}(\mathcal{O}^K)$. In particular, Eve has an almost-surely winning observation-based strategy in \mathbb{G} if and only if she has one in \mathbb{G}^K .

In the setting of the previous proposition, consider the special case where Eve has an almost-surely winning observation-based strategy φ^K in \mathbb{G}^K that only depends on the current knowledge (in particular, it is *memoryless*). Then the corresponding almost-surely winning observation-based strategy φ in \mathbb{G} is, in general, not memoryless, but can be implemented by a finite transducer whose set of control states is precisely the set of possible knowledges for Eve. More precisely the strategy consists in computing and updating on the fly (using a finite automaton) the value of the knowledge after the current partial play and to pick the next action by solely considering the knowledge. We may refer at such a strategy φ as a **knowledge-only strategy**.

4 Reachability Objectives

The main result of this section is the following.

Theorem 1. *For any reachability concurrent game with imperfect information, one can decide, in doubly exponential time, whether Eve has an almost-surely winning strategy. If Eve has such a strategy then she has a knowledge-only uniform strategy, and such a strategy can be effectively constructed.*

Before proving Theorem 1 we first establish an intermediate result on positively winning in $1\frac{1}{2}$ -player safety game

4.1 Positively winning in $1\frac{1}{2}$ -player safety game with imperfect information

A concurrent game (with imperfect information) in which one player has only a single available action is what we refer as a **$1\frac{1}{2}$ -player game with imperfect information** (those games are also known in the literature as *partially observable Markov Decision Processes*). The following result is a key ingredient for the proofs of Proposition 3 and Theorem 1.

Lemma 1. *Consider an $1\frac{1}{2}$ -player safety game with imperfect information. Assume that the player has an observation-based strategy that is positively winning. Then she also has an observation-based finite memory strategy that is positively winning. Moreover, both the strategy and the set of positively winning states can be computed in time $\mathcal{O}(2^{|S|})$.*

We first simplify the notations for the special case of $1\frac{1}{2}$ -player games. We define an **$1\frac{1}{2}$ -player arena** as a tuple $\mathcal{A} = \langle S, \Sigma, \delta, F \rangle$ where S is a finite set of control states, Σ is a finite set of actions, $F \subseteq S$ is a set of final states and $\delta : S \times \Sigma \rightarrow \mathcal{D}(S)$ is the transition (total) function. A **$1\frac{1}{2}$ -player game** is a tuple $\mathbb{G} = (\mathcal{A}, \sim, s_0, \mathcal{O})$ where \sim is an equivalence relation on states, $s_0 \in S$ is an initial state and \mathcal{O} is an objective (safety in the sequel). The notions of knowledge, knowledge arena and knowledge-based (memoryless) strategy are trivially adapted to this setting.

We will be interested in a special kind of strategies, that we call **ultimately knowledge-based memoryless strategy**. Such a strategy consists in playing randomly trying to reach a target state t and then play in a knowledge-based memoryless fashion assuming that t as been effectively reached. More formally, a strategy φ is an **ultimately knowledge-based memoryless strategy** if it is of the following form (where $t \in S$ and $k < |S|$): play uniformly randomly any action in Σ on the k first moves and then set knowledge to be $\{t\}$, update it along the play and pick the moves only depending on the current knowledge. Hence, it is of the following form:

- $\varphi(s_0 \sigma_{E,0} s_1 \cdots s_n) = d_{\text{univ}}$ if $n < k$;
- otherwise, $\varphi(s_0 \sigma_{E,0} s_1 \cdots s_n) = d_{\text{Knowledge}(t \sigma_{E,k} \cdot s_{k+1} \sigma_{E,k+1} \cdots s_n)}$ where d_{univ} is the uniform distribution over Σ and $d_K \in \mathcal{D}(\Sigma)$ for any $K \subseteq S$.

In particular those strategies are *finite memory strategy* (and the memory needed is the set $\{1, \dots, k\} \cup 2^S$).

Lemma 1 is a direct consequence of the following slightly more precise lemma.

Lemma 2. *Consider a $1\frac{1}{2}$ -player safety game with imperfect information. Assume that the player has an observation-based strategy that is positively winning. Then she also has an ultimately knowledge-based memoryless strategy that is positively winning. Moreover, both the strategy and the set of positively winning states can be computed in time $\mathcal{O}(2^{|S|})$.*

Proof. In the sequel we will be interested in computing the set $S^{>0}$ of those states $s \in S$ for which the player has an observation-based strategy that is positively winning.

Consider the knowledge arena \mathcal{A}^K and the knowledge game \mathbb{G}^K associated with \mathcal{A} and \mathbb{G} . We define a notion of almost-surely winning knowledge by letting

$$\begin{aligned} \mathcal{K}^{\text{AS}} = \{K \in 2^S \mid \exists \varphi \text{ knowledge-based strategy s.t. } \forall s \in K, \\ \varphi \text{ is almost-surely winning for the player in } \mathbb{G}^K \text{ from } (s, K)\} \end{aligned}$$

We claim that a knowledge $K \in \mathcal{K}^{\text{AS}}$ is actually surely winning: there exists a knowledge-based strategy φ such that $\forall s \in K, \text{Outcomes}(\varphi, (s, K)) \subseteq ((S \setminus F)\Sigma_E)^\omega$, i.e. playing accordingly to φ the player is sure that no play will be loosing for her. Indeed, consider the following (decreasing and bounded) sequence of knowledges:

$$\begin{cases} \mathcal{K}_0 = 2^{S \setminus F} \\ \mathcal{K}_{i+1} = \mathcal{K}_i \cap \text{Pre}(\mathcal{K}_i) \end{cases}$$

where

$$\text{Pre}(\mathcal{K}) = \{K \in 2^{S \setminus F} \mid \exists \sigma, \forall s \in K, \delta^K((s, K), \sigma)((s', K')) > 0 \Rightarrow K' \in \mathcal{K}\}$$

is the set of (non-final) knowledges from which the player is sure that in the next step the play will be in \mathcal{K} .

Let \mathcal{K}^* be the limit of the sequence $(\mathcal{K}_i)_{i \geq 0}$. Then, we have the following fact.

Fact 1 *The following equality holds: $\mathcal{K}^{\text{AS}} = \mathcal{K}^*$.*

Proof. The inclusion $\mathcal{K}^{\text{AS}} \supseteq \mathcal{K}^*$ is immediate and it also come with a deterministic surely winning strategy for the player that simply consists in playing an action that ensures to stay inside \mathcal{K}^* (such an action exists by definition of the *Pre* operator). For the converse inclusion, we prove that a knowledge $K \notin \mathcal{K}^*$ cannot be in \mathcal{K}^{AS} . For such a K we define its rank $rk(K) = i$ to be unique interger i such that $K \in \mathcal{K}_i \setminus \mathcal{K}_{i+1}$ (by convention we let $\mathcal{K}_{-1} = 2^S$) and we prove the result by induction on $rk(K)$. For $rk(K) = 0$, the result is immediate. Now, assume that it holds for some i and let K be some knowledge with $rk(K) = i + 1$: for all $\sigma \in \Sigma$, there is some $s \in S$ and some configuration (s', K') with $\delta^K((s, K), \sigma)((s', K')) > 0$ and $rk(K') \leq i$: hence for any knowledge-based strategy φ there is some $s \in K$ such that the probability, after one move starting from (s, K) , of reaching a configuration with a knowledge having a rank smaller or equal than i , is strictly positive. One concludes then by induction that $K \notin \mathcal{K}^{\text{AS}}$. $\square_{[\text{Fact 1}]}$

The following short example gives some intuition for the next construction.

Example 1. Consider the arena depicted in Figure 1 (edges are labeled by both the pair of actions and the probability of reaching their target; final states are double circled). We assume that in this example $s \sim t \sim t'$.

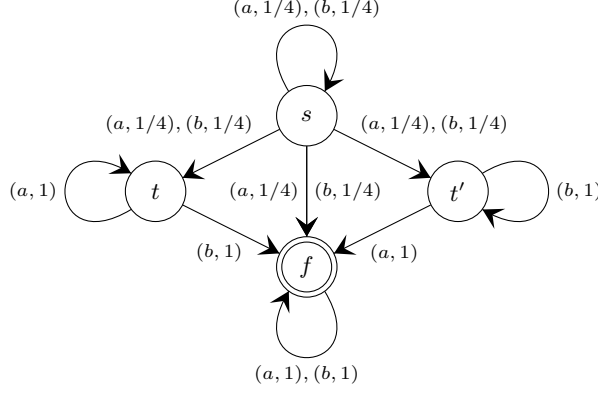


Fig. 1. Arena of Example 1

In this example $\mathcal{K}^{\text{AS}} = \{\{t\}, \{t'\}\}$. Note also that if one considers the corresponding knowledge game, the configurations $(t, \{t\})$ and $(t', \{t'\})$ are never visited in a play starting from $(s, \{s\})$. Nevertheless the player has a strategy that wins with probability $1/4$ starting from state s : the first action is to play randomly (with same probability) action a or b , and then to play a forever. The intuitive idea behind this strategy is rather simple: there are two safe states, t and t' , that is states from which the player surely wins. Moreover those states can be possibly reached from s (they belong to the same connected component) but this cannot be detected due to the equivalence relation \sim . Hence the player bets that one of those states is reached (here state t) and behaves like if it is the case. Playing randomly on her first move is a way to ensure that with some positive probability (here $1/4$) state t is reached. As from t the player can surely win, following the associated (knowledge-based memoryless) strategy, she is sure to win if her bet was correct: altogether it provides a strategy that is winning with probability $1/4$. $\square_{\text{[Example 1]}}$

Define $S^{\text{AS}} = \{s \in S \mid \{s\} \in \mathcal{K}^{\text{AS}}\}$ and consider the following (increasing and bounded) sequence:

$$\begin{cases} W_0 = S^{\text{AS}} \\ W_{i+1} = W_i \cup \{s \in S \setminus F \mid \exists \sigma \in \Sigma \text{ and } t \in W_i \text{ s.t. } \delta(s, \sigma)(t) > 0\} \end{cases}$$

Let W be the limit of the sequence $(W_i)_{i \geq 0}$: it consists exactly of those states from which the player has a strategy ensuring to reach, in at most $|S|$ moves

and without going through a final state, a state in S^{AS} with some positive probability. Also note that the corresponding strategy is the one that plays with equal probability any action in Σ .

The states in W are actually those from which the player has a positively winning strategy.

Fact 2 *The following equality holds: $W = S^{>0}$.*

Proof. The inclusion $W \subseteq S^{>0}$ is rather immediate: from some state $s \in W$, the player should first play randomly on the first k rounds (where k is the smallest integer such that $s \in W_k$ in the previously defined sequence) and then play as if it was in some state $t \in S^{\text{AS}}$ (where t is some reachable state from s in k moves accordingly to the definition of the sequence $(W_i)_{i \geq 0}$). This last step is done using an observation-based strategy that mimics the one coming with the construction of \mathcal{K}^{AS} : the only trick here is that the player "reset" the knowledge to be $\{t\}$.

Consider now the converse inclusion: $S^{>0} \subseteq W$. Let s be some state in $S^{>0}$ and assume, by contradiction, that $s \notin W$. In particular a play starting from s will never go through a state in S^{AS} . We claim that playing accordingly to some knowledge-based strategy is almost-surely loosing for the player: indeed, as a consequence of the definition of S^{AS} (and of \mathcal{K}^{AS}), from a state $t \notin S^{\text{AS}}$, playing a knowledge-based strategy, the probability of visiting a final state in the next $2^{|S|}$ moves is some $\varepsilon > 0$. Moreover such a play stays outside of S^{AS} forever. Hence, using Borel-Cantelli Lemma, the probability that such a play never goes through a final state is 0, meaning that it is almost-surely loosing, thus contradicting the assumption that $s \in S^{>0}$. Hence, $s \in W$ which concludes the proof. $\square_{[\text{Fact 2}]}$

Now one can conclude the proof of Lemma 1: the existence of the ultimately knowledge-based memoryless strategy follows from Fact 2. Effectivity as well as complexity come from the constructive way of defining W (and other intermediate objects) by means of fixpoint computations. $\square_{[\text{Lemma 2}]}$

4.2 Proof of Theorem 1

Fix, for the rest of this section, a concurrent game with imperfect information $\mathbb{G} = (\mathcal{A}, s_0, \mathcal{O})$ equipped with a reachability objective \mathcal{O} defined from a set F of final states. We set $\mathcal{A} = \langle S, \Sigma_E, \Sigma_A, \delta, \sim_E, \sim_A \rangle$. We also consider $\mathbb{G}^K = (\mathcal{A}^K, (s_0, \{s_0\}), \mathcal{O}^K)$ to be the corresponding knowledge game.

To prove Theorem 1, one first defines (in a non constructive way) a knowledge-only uniform strategy φ for Eve as follows. We let

$$\mathcal{K}^{\text{AS}} = \{K \in 2^S \mid \exists \varphi_E \text{ knowledge-based strategy for Eve s.t. } \varphi_E \text{ is almost-surely winning for Eve in } \mathbb{G}^K \text{ from any } (s, K) \text{ with } s \in K\}$$

be the set of knowledges made only by almost-surely winning states for Eve (note here that we require that the almost-surely winning strategy is the same *for all* states with the same knowledge).

For every knowledge $K \in \mathcal{K}^{\text{AS}}$ we define

$$\text{Allow}(K) = \{\sigma_E \in \Sigma_E \mid \forall s \in K, \forall \sigma_A \in \Sigma_A, \delta^K((s, K), \sigma_E, \sigma_A)((s', K)) > 0 \Rightarrow K' \in \mathcal{K}^{\text{AS}}\}$$

Proposition 2. *For every knowledge $K \in \mathcal{K}^{\text{AS}}$, $\text{Allow}(K) \neq \emptyset$.*

Proof. Consider some knowledge $K \in \mathcal{K}^{\text{AS}}$ and assume by contradiction that $\text{Allow}(K) = \emptyset$. As $K \in \mathcal{K}^{\text{AS}}$, there exists, by definition, some knowledge-based strategy φ_E for Eve in \mathbb{G}^K that is almost-surely winning from any state (s, K) with $s \in K$. Strategy φ_E gives the same distribution for all partial plays of the form (s, K) with $s \in K$ (i.e. consisting of a single state indistinguishable by Eve). Moreover there exists some action $\sigma_E \in \Sigma_E$ such that $\varphi_E(s, K)(\sigma_E) > 0$ and, as $\text{Allow}(K) = \emptyset$, there is some action $\sigma_A \in \Sigma_A$ and some configuration (s', K') such that $\delta^K((s, K), \sigma_E, \sigma_A)(s', K') > 0$ and $K' \notin \mathcal{K}^{\text{AS}}$. Consider the strategy φ'_E for Eve defined by $\varphi'_E(\lambda) = \varphi_E((s, K) \cdot \sigma_E \cdot \lambda)$, i.e. φ'_E mimics strategy φ_E . In particular, as $K' \notin \mathcal{K}^{\text{AS}}$, it means that there is some $s'' \in K'$ such that φ'_E is not almost-surely winning for Eve in \mathbb{G}^K from (s'', K') . Now, by considering the way the knowledge is updated, one concludes that there is some state $s''' \in S$ such that $\delta^K((s''', K), \sigma_E, \sigma_A)(s'', K') > 0$ (as s'' belongs to a knowledge updated from K). Now, if one considers the strategy of Adam that plays σ_A and then mimics a counter strategy against φ'_E , it follows that this strategy ensures Adam to win with positive probability while playing against φ_E starting from (s''', K) . This contradicts our initial assumption on φ_E being almost-surely winning from all configuration of the form (s, K) with $s \in K$. Hence $\text{Allow}(K) \neq \emptyset$. \square

We now consider a (well-defined) knowledge-based uniform memoryless strategy φ for Eve on $K \in \mathcal{K}^{\text{AS}}$ by letting

$$\varphi(K)(\sigma_E) = \begin{cases} \frac{1}{|\text{Allow}(K)|} & \text{if } \sigma_E \in \text{Allow}(K) \\ 0 & \text{otherwise} \end{cases}$$

The next proposition shows that φ is almost-surely winning for Eve.

Proposition 3. *The strategy φ is almost-surely winning for Eve from states whose Eve's knowledge is in \mathcal{K}^{AS} .*

Proof. In order to prove that φ is almost-surely winning one needs to show that it is almost-surely winning against *any* observation-based strategy φ_A of Adam, i.e. $\text{Pr}_{(s_0, \{s_0\})}^{\varphi, \varphi_A}(\mathcal{O}) = 1$. Note that once φ is fixed, and as it is a knowledge-based memoryless strategy, it induces a new game, denoted \mathbb{G}_φ , where only Adam makes choices, and where φ 's choices are now part of the stochastic aspect of the game. More formally, \mathbb{G}_φ is the $1\frac{1}{2}$ -player *safety* game $\mathbb{G}_\varphi = (\mathcal{A}_\varphi, \sim_A^K, (s_0, \{s_0\}), \mathcal{O})$ where $\mathcal{A}_\varphi = \langle S^K, \Sigma_A, \delta_\varphi, F^K \rangle$ with

$$- \delta_\varphi((s, K), \sigma_A)(s', K') = \sum_{\sigma_E \in \Sigma_E} \delta^K((s, K), \sigma_E, \sigma_A)(s', K') \cdot \varphi(K)(\sigma_E)$$

As a strategy for Adam in \mathbb{G} (equivalently in \mathbb{G}^K) can be seen as well as a strategy in game \mathbb{G}_φ and *vice versa*, while preserving the value of the game (against φ in \mathbb{G}), one derives the following fact.

Fact 3 *Strategy φ is almost-surely winning in \mathbb{G}^K if and only if the player has no positively winning observation-based strategy in \mathbb{G}_φ .*

As \mathbb{G}_φ is a $1\frac{1}{2}$ -player safety concurrent game with imperfect information, one can use the previous results. In particular Lemma 2 implies that in order to prove Proposition 3, it suffices to prove that φ is winning against any strategy in \mathbb{G}^K that is obtained by mimicking an ultimately memoryless knowledge-based strategy of the player in \mathbb{G}_φ ¹. As those strategies are finite memory strategies one derives the following fact.

Fact 4 *Strategy φ is almost-surely winning in \mathbb{G}^K if and only if it is almost-surely winning against any finite-memory observation-based strategy of Adam.*

Fix such a finite memory strategy $\psi = (\text{Move}, \text{Up}, m_0)$ for Adam (let M be the finite memory used here). Now from M , Up and \mathcal{A}^K one can construct a new arena whose set of states is $S^K \times M$ and where the M component is updated accordingly to Up : the idea is just to make explicit the value of the memory at any stage and to update it explicitly in the arena. Next, if one modifies the equivalence relation of Adam \sim_A to only distinguish between those configurations that have a different memory content, and if one modifies \sim_E so that Eve has no information on the M component, one obtains a new game in which any finite-memory strategy of Adam with memory M and update function Up in the previous game is transformed into a memoryless observation-based strategy for him.

More formally, this leads to consider the arena

$$\mathcal{A}_{\text{Up}}^K = \langle S^K \times M, \Sigma_E, \Sigma_A, \delta_{\text{Up}}^K, F^K \times M \rangle \text{ where}$$

$$\delta_{\text{Up}}^K((s, K, m), \sigma_E, \sigma_A)(s', K', m') = 0 \text{ if } m' \neq \text{Up}(m, [s']_{/\sim_A}, \sigma_A) \text{ and}$$

$$\delta_{\text{Up}}^K((s, K, m), \sigma_E, \sigma_A)(s', K', m') = \delta^K((s, K), \sigma_E, \sigma_A)(s', K') \text{ otherwise.}$$

The new equivalence relations are given by $(s, K, m) \equiv_E (s', K', m')$ iff $(s, K) \sim_E^K (s', K')$ and $(s, K, m) \equiv_A (s', K', m')$ iff $m = m'$. Let \mathbb{G}_{Up}^K be this

¹ An ultimately memoryless knowledge-based strategy of the player in \mathbb{G}_φ is a finite memory strategy that uses as a memory a set of integers $\{1, \dots, k\}$ (for the initial part) together with a set of knowledge. Note that a knowledge \mathbb{G}_φ is a subset of states, *i.e.* is a subset of pairs made of a control state $s \in S$ and of a subset $K \subseteq S$ that represents a knowledge of Eve in the previous game \mathbb{G} (hence Adam is not only computing the possible states he can be in but he is also computing the set of knowledges Eve can have about the play. This somehow proves that to positively wins against φ he would use doubly exponential size memory). Nevertheless the only important thing is that it is a finite memory strategy, and that it can hence be translated in another finite memory strategy in \mathbb{G}^K .

new game. Any strategy for Eve in \mathbb{G}^K can be seen as a strategy in \mathbb{G}_{Up}^K and vice versa.

Hence from now on we may only work in \mathbb{G}_{Up}^K and assume that ψ is a memoryless knowledge-based strategy for Adam and our goal is to prove that φ almost-surely wins against ψ from any configuration in $\{(s, K, m_0) \mid K \in \mathcal{K}^{\text{AS}} \text{ and } s \in K\}$. Actually, one will prove a slightly stronger result, namely that φ almost-surely wins against ψ from any configuration in $\{(s, K, m) \mid m \in M, K \in \mathcal{K}^{\text{AS}} \text{ and } s \in K\}$

We will first define an increasing sequence of subsets of almost-surely winning positions for Eve in \mathbb{G}_{Up}^K and later we will prove that its limit is the set of all positions with a knowledge in \mathcal{K}^{AS} and that φ is actually an almost-surely winning strategy from those positions. For some configuration (s, K, m) and some action $\sigma_E \in \Sigma_E$ and some distribution of actions d in $\mathcal{D}(\Sigma_A)$, we define $\text{Post}_{\sigma_E, d}((s, K, m))$ as the set of all possible next states when Eve plays σ_E and Adam picks an action according to d from (s, K, m) :

$$\begin{aligned} \text{Post}_{\sigma_E, d}((s, K, m)) &= \{(s', K', m') \mid \exists \sigma_A \text{ s.t. } d(\sigma_A) > 0 \\ &\quad \text{and } \delta^K((s, K, m), \sigma_E, \sigma_A)((s', K', m')) > 0\} \end{aligned}$$

Consider the following increasing sequence $(\text{Rank}_i)_{i \geq 0}$: $\text{Rank}_0 = 2^{F^K} \times M$ consists of trivially winning positions and

$$\begin{aligned} \text{Rank}_{i+1} &= \text{Rank}_i \cup \{(s, K, m) \mid K \in \mathcal{K}^{\text{AS}} \\ &\quad \text{and } \exists \sigma_E \in \text{Allow}(K) \text{ s.t. } \text{Post}_{\sigma_E, \text{Move}(m)}(s, K, m) \cap \text{Rank}_i \neq \emptyset\} \end{aligned}$$

Let us denote by Rank^* the limit of the sequence $(\text{Rank}_i)_{i \geq 0}$. We claim that $\text{Rank}^* = \{(s, K, m) \mid K \in \mathcal{K}^{\text{AS}}\}$ and that φ is an almost-surely winning strategy for Eve from those positions when Adam plays accordingly to ψ .

The inclusion $\text{Rank}^* \subseteq \{(s, K, m) \mid K \in \mathcal{K}^{\text{AS}}\}$ is forced by the definition of Rank^* . The fact that φ is an almost-surely winning strategy for Eve from positions in Rank^* when Adam plays accordingly to ψ is a simple consequence of how $(\text{Rank}_i)_{i \geq 0}$ is defined and of Borel-Cantelli Lemma. Indeed from any configuration in Rank_i , there is a non null probability to reach a final state in the next i moves while playing φ against ψ and moreover the play surely stays inside Rank^* while playing φ against ψ (at least until some final state is visited): hence the probability, that a play starting from Rank^* , in which Eve follows φ and Adam follows ψ , to never reach $F^K \times M$ is null.

In order to prove the other inclusion, we let $X = \{(s, K, m) \mid K \in \mathcal{K}^{\text{AS}}\} \setminus \text{Rank}^*$ and assume by contradiction that $X \neq \emptyset$. By definition, for any element $(s, K, m) \in X$ we have that $\forall \sigma_E \in \text{Allow}(K)$, $\text{Post}_{\sigma_E, \text{Move}(m)}(s, K, m) \subseteq X$. This means in particular that following φ from such a configuration, and if Adam plays accordingly to ψ , then Eve surely loses as the play surely stays in X and $X \cap 2^{F^K} \times M = \emptyset$. Now we claim that the same holds if one replaces φ by any almost-surely winning strategy for Eve, leading to a contradiction. Indeed consider an almost-surely winning strategy φ^{AS} for Eve. Then we claim the following fact.

Fact 5 *Let λ be a partial play consisting only of configurations in X . Assume that, for some strategy ψ' of Adam, λ is a possible partial play accordingly to both φ^{AS} and ψ' (more formally, $\Pr_{(s,K,m)}^{\varphi,\psi'}(\text{cone}(\lambda)) > 0$ where (s,K,m) denotes the initial configuration of λ). Then for any action $\sigma_E \in \Sigma_E$, one has $\varphi^{\text{AS}}(\lambda)(\sigma_E) > 0$ only if $\sigma_E \in \text{Allow}(K')$ where K' denotes Eve's knowledge in the last configuration of λ .*

Proof. The proof is by contradiction. Consider some $\lambda, \varphi^{\text{AS}}, \psi, \sigma_E$ violating the property. Then one can find an equivalent play $\lambda' \sim_E \lambda$ such that λ' ends in a configuration (s', K', m') and there is an action $\sigma_A \in \Sigma_A$ such that $\delta^K((s', K', m'), \sigma_E, \sigma_A)((s'', K'', m'')) > 0$ for some $K'' \notin \mathcal{K}^{\text{AS}}$ (existence of λ' follows by the construction of the knowledge arena). Now consider the strategy of Adam that first mimics ψ and then after $|\lambda'|$ moves plays σ_A and then plays accordingly to a strategy ensuring from (s'', K'', m'') that Eve's does not surely wins (such a strategy exists as $K'' \notin \mathcal{K}^{\text{AS}}$): then against this strategy φ^{AS} is not almost-surely winning, leading a contradiction $\square_{[\text{Fact5}]}$

Now one is ready to conclude. Assume Adam plays accordingly to ψ and Eve plays accordingly to some almost-surely winning strategy φ^{AS} . Then it follows from Fact 5 and definition of X that a play starting in X stays forever in X , hence never visits $F^K \times M$ and contradicting the hypothesis that φ^{AS} is almost-surely winning. Therefore $X = \emptyset$, which concludes the proof of Proposition 3. \square

Now one can prove Theorem 1. First Eve almost-surely wins in \mathbb{G} if and only if she almost-surely wins in \mathbb{G}^K if and only if $\{s_0\} \in \mathcal{K}^{\text{AS}}$, *i.e.* (using Proposition 3) if and only if Eve has a knowledge-only uniform strategy in \mathbb{G}^K . Now, to decide whether Eve almost-surely wins in \mathbb{G} , it suffices to check, for any possible knowledge-only uniform strategy φ for her, whether it is almost-surely winning. Once φ is fixed, it leads, from Adam's point of view, to a $1\frac{1}{2}$ -player safety game \mathbb{G}_φ where the player positively wins if and only if φ is not almost-surely winning. Hence Lemma 1 implies that deciding whether φ is almost-surely winning can be done in time exponential in the size of \mathbb{G}_φ , which itself is of exponential size in $|S|$. Hence deciding whether a knowledge-only uniform strategy for Eve is winning can be done in doubly exponential time (in the size of $|S|$). The set of knowledge-only uniform strategies for Eve is finite and its size is doubly exponential in the size of the game. Hence the overall procedure, that tests every possible such strategies, requires doubly exponential time. As effectivity is immediate, this concludes the proof of Theorem 1.

4.3 Complexity Lower Bound

The naive underlying algorithm of Theorem 1 turns out to be optimal.

Theorem 2. *Deciding whether Eve almost-surely wins a concurrent game with imperfect information is a 2-EXPTIME-complete problem.*

Proof (sketch). The proof is a generalisation of a similar result given in [8] showing EXPTIME-hardness of concurrent games *only one* player is imperfectly informed. The idea is to simulate an alternating exponential space Turing machine (without input). We design a game where the players describe the run of such a machine: transitions from existential (*resp.* universal) states are chosen by Eve (*resp.* Adam) and Adam is also in charge of describing the successive configurations of the machine. To prevent him from cheating, Eve can secretly mark a cell of the tape, and latter check whether it was correctly updated (if not she wins). As she cannot store the exact index of the cell (it is of exponential size), she could cheat in the previous phase: hence Adam secretly marks some bit and one recall the value of the corresponding bit of the index of the marked cell: this bit is checked when Eve claims that Adam cheated (if it is wrong then she is loosing). Eve also wins if the described run is accepting. Eve can also restart the computation whenever she wants (this is useful when she cannot prove that Adam cheated): hence if the machine accepts the only option for Adam is to cheat, and Eve will eventually catch him with probability one. Now if the machine does not accept, the only option for Eve is to cheat, but it will be detected with positive probability. \square

5 Büchi Objectives

We now consider the problem of deciding whether Eve almost-surely wins a Büchi game and establish the following result.

Theorem 3. *For any Büchi concurrent game with imperfect information, one can decide, in doubly exponential time, whether Eve has an almost-surely winning strategy. If Eve has such a strategy then she has a knowledge-based uniform memoryless strategy, and such a strategy can be effectively constructed. The doubly exponential time complexity bound is optimal.*

The results and techniques are similar to the one for reachability games. In particular, we need first to establish an intermediate result about positive winning in $1\frac{1}{2}$ -player co-Büchi game with imperfect information.

5.1 Positively winning in $1\frac{1}{2}$ -player co-Büchi game with imperfect information

We have the following lemma whose statement and proof are very similar to the one of Lemma 1 except that now the winning states are those connected by *any* kind of path to a surely winning state).

Lemma 3. *Consider an $1\frac{1}{2}$ -player co-Büchi game with imperfect information. Assume that the player has an observation-based strategy that is positively winning. Then she also has an observation-based finite memory strategy that is positively winning. Moreover, both the strategy and the set of positively winning states can be computed in time $\mathcal{O}(2^{|S|})$.*

Lemma 3 is a direct consequence of the following slightly more precise lemma.

Lemma 4. *Consider an $1\frac{1}{2}$ -player co-Büchi game with imperfect information. Assume that the player has an observation-based strategy that is positively winning. Then she also has an ultimately knowledge-based memoryless strategy that is positively winning. Moreover, both the strategy and the set of positively winning states can be computed in time $\mathcal{O}(2^{|S|})$.*

Proof. Fix an arena $\mathcal{A} = \langle S, \Sigma, \delta \rangle$ together with a co-Büchi game $\mathbb{G} = (\mathcal{A}, \sim, s_0, F)$ (recall here that a play is winning if and only if it visits finitely often the set F).

In the sequel we will be interested in computing the set $S^{>0}$ of those states $s \in S$ from which the player has an observation-based strategy that is positively winning.

The proof is very similar to the one of Lemma 2, and therefore we reuse the notations as well as intermediate results. Recall that to win a safety game with positive probability the player needs to reach a state from which she is sure not to visit a final state (those states were denoted S^{AS}). In order to reach such a state, the player plays randomly on some fixed number of initial moves. Then she bets that she reached some good state $s \in S^{\text{AS}}$ and plays as if it is the case and in such a way to surely win (actually the state s is guessed from the very beginning and is part of the strategy). As this second stage can be done mimicking a knowledge-based memoryless strategy, this leads to an ultimately knowledge-based memoryless strategy.

An important point is that the state s should be reached without visiting a final state in the meantime. If the game is equipped now with a co-Büchi objective, it no longer matter. This remark leads to the following definition. Consider the following (increasing and bounded) sequence:

$$\begin{cases} \hat{W}_0 = S^{\text{AS}} \\ \hat{W}_{i+1} = \hat{W}_i \cup \{s \in S \mid \exists \sigma \in \Sigma \text{ and } t \in \hat{W}_i \text{ s.t. } \delta(s, \sigma)(t) > 0\} \end{cases}$$

Let \hat{W} be the limit of the sequence $(\hat{W}_i)_{i \geq 0}$: it consists exactly of those states from which the player has a strategy ensuring her to reach, in at most $|S|$ moves, a state in S^{AS} with some positive probability. Also note that the corresponding strategy is the one that plays randomly (with equiprobability) any action in Σ . Hence the only difference between the set \hat{W} and the set W defined in the proof of Lemma 2 is that one allows to have final states in the fixpoint definition of \hat{W} .

The states in \hat{W} are actually those from which the player can win the co-Büchi game with a positive probability.

Fact 6 *The following equality holds: $\hat{W} = S^{>0}$.*

Proof. The inclusion $\hat{W} \subseteq S^{>0}$ is rather immediate: from some state $s \in \hat{W}$, the player should first play randomly on the first i rounds (where i is the smallest integer such that $t \in \hat{W}_i$ in the previously defined sequence) and then play as

if it was in some state $t \in S^{\text{AS}}$ (where t is some reachable state from s in i moves accordingly to the definition of the sequence $(\hat{W}_i)_{i \geq 0}$). This last step is done using an observation-based strategy that mimics the one coming with the construction of \mathcal{K}^{AS} : the only trick here is that the player "reset" the knowledge to be $\{t\}$.

Consider now the converse inclusion: $S^{>0} \subseteq \hat{W}$. Let s be some state in $S^{>0}$ and assume, by contradiction, that $s \notin \hat{W}$. In particular a play starting from s will never go through a state in S^{AS} . We claim that playing accordingly to some knowledge-based strategy is almost-surely loosing for the player: indeed, as a consequence of the definition of S^{AS} (and of \mathcal{K}^{AS}), from a state $t \notin S^{\text{AS}}$, playing a knowledge-based strategy, the probability of visiting a final state in the next $2^{|S|}$ moves is some $\varepsilon > 0$. Moreover such a play stays outside of S^{AS} forever. Hence, for any $k > 0$, using Borel-Cantelli Lemma, the probability of such a play to visit at most k final states is 0 hence implying that the probability of going finitely often through F is 0 too. This means that s is surely loosing for the player and contradicts the initial hypothesis of $s \in S^{>0}$. Hence, $s \in \hat{W}$ which concludes the proof. $\square_{[\text{Fact } 6]}$

Now one can conclude the proof of Lemma 3: the existence of the ultimately knowledge-based memoryless strategy follows from Fact 2. Effectivity as well as complexity comes from the constructive way of defining \hat{W} (and other intermediate objects) by means of fixpoint computations. $\square_{[\text{Lemma } 3]}$

5.2 Proof of Theorem 3

We now turn to the proof of Theorem 3. Again, the key idea is to prove that the strategy that plays randomly inside the almost-surely winning region is an almost-surely winning strategy.

For the rest of this section, we fix a concurrent game with imperfect information $\mathbb{G} = (\mathcal{A}, \sim_E, \sim_A, s_0, \mathcal{O})$ equipped with a Büchi objective \mathcal{O} . We set $\mathcal{A} = \langle S, \Sigma_E, \Sigma_A, \delta, F \rangle$. We also consider $\mathbb{G}^K = (\mathcal{A}^K, \sim_E^K, \sim_A^K, (s_0, \{s_0\}), \mathcal{O}^K)$ to be the corresponding knowledge game.

The proof follows the same line as the one to prove Theorem 1. The main idea is again to prove that the strategy that stay inside the configuration with an almost-surely winning knowledge is almost-surely winning.

Again, we let

$$\begin{aligned} \mathcal{K}^{\text{AS}} = \{ & K \in 2^S \mid \exists \varphi_E \text{ knowledge-based strategy for Eve s.t. } \varphi_E \text{ is} \\ & \text{almost-surely winning for Eve in } \mathbb{G}^K \text{ from any } (s, K) \text{ with } s \in K \} \end{aligned}$$

be the set of equivalent classes (with respect to \sim_E^K) made only by almost-surely winning states for Eve (note here that we require that the almost-surely winning strategy is the same for all configurations with the same knowledge). For every knowledge $K \in \mathcal{K}^{\text{AS}}$ we define

$$\begin{aligned} \text{Allow}(K) = \{ & \sigma_E \in \Sigma_E \mid \forall s \in K, \forall \sigma_A \in \Sigma_A, \\ & \delta^K((s, K), \sigma_E, \sigma_A)((s', K)) > 0 \Rightarrow K' \in \mathcal{K}^{\text{AS}} \} \end{aligned}$$

We then get a result similar to Proposition 2 (the proof is exactly the same as the one of Proposition 2).

Proposition 4. *For every knowledge $K_E \in \mathcal{K}_E^{\text{AS}}$, $\text{Allow}(K_E) \neq \emptyset$.*

Again, we define now a knowledge-based uniform memoryless strategy φ for Eve on $K \in \mathcal{K}^{\text{AS}}$ by letting

$$\varphi(K)(\sigma_E) = \begin{cases} \frac{1}{|\text{Allow}(K)|} & \text{if } \sigma_E \in \text{Allow}(K) \\ 0 & \text{otherwise} \end{cases}$$

The next proposition shows that φ is almost-surely winning for Eve.

Proposition 5. *The strategy φ is almost-surely winning for Eve from states whose Eve's knowledge is in $\mathcal{K}_E^{\text{AS}}$.*

Proof. In order to prove that φ is almost-surely winning one needs to show that it is almost-surely winning against *any* observation-based strategy φ_A of Adam, i.e. $\text{Pr}_{(s_0, \{s_0\})}^{\varphi, \varphi_A}(\mathcal{O}) = 1$. Again, as in the reachability case, once φ is fixed, and as it is a knowledge-based memoryless strategy, it induces a $1\frac{1}{2}$ -player co-Büchi game denoted \mathbb{G}_φ and defined exactly as in the proof of Proposition 3.

As a strategy for Adam in \mathbb{G} (equivalently in \mathbb{G}^K) can be seen as well as a strategy in game \mathbb{G}_φ and *vice versa*, while preserving the value of the game (against φ in \mathbb{G}), one derives the following fact.

Fact 7 *Strategy φ is almost-surely winning in \mathbb{G}^K if and only if the player has no positively winning observation-based strategy in the co-Buchi \mathbb{G}_φ .*

As \mathbb{G}_φ is a $1\frac{1}{2}$ -player co-Büchi concurrent game with imperfect information, one can use Lemma 4 to conclude that, in order to prove Proposition 3, it suffices to prove that φ is winning against any *finite-memory* observation-based strategy of Adam in \mathbb{G}^K .

Fact 8 *Strategy φ is almost-surely winning in \mathbb{G}^K if and only if it is almost-surely winning against any finite-memory observation-based strategy of Adam.*

Fix such a finite memory strategy $\psi = (\text{Move}, \text{Up}, m_0)$ for Adam (let M be the finite memory used here). As in the proof of Proposition 3 it leads to define a new game \mathbb{G}_{Up}^K (we keep here the same notations / definitions). Again, we may assume that ψ is a memoryless knowledge-based strategy for Adam and our goal is to prove that φ almost-surely wins against ψ from any configuration in $\{(s, K, m_0) \mid K \in \mathcal{K}^{\text{AS}} \text{ and } s \in K\}$. Actually, one will prove a slightly stronger result, namely that φ almost-surely wins against ψ from any configuration in $\{(s, K, m) \mid m \in M, K \in \mathcal{K}^{\text{AS}} \text{ and } s \in K\}$.

We will first define an increasing sequence of subsets of almost-surely winning positions for Eve in \mathbb{G}_{Up}^K and later we will prove that its limit is the set of all positions with a knowledge in \mathcal{K}^{AS} and that φ is actually an almost-surely

winning strategy from those positions. For some configuration (s, K, m) and some action $\sigma_E \in \Sigma_E$ and some distribution of actions d in $\mathcal{D}(\Sigma_A)$, we define $\text{Post}_{\sigma_E, d}((s, K, m))$ as the set of all possible next states when Eve plays σ_E and Adam picks an action according to d from (s, K, m) :

$$\begin{aligned} \text{Post}_{\sigma_E, d}((s, K, m)) = \{ & (s', K', m') \mid \exists \sigma_A \text{ s.t. } d(\sigma_A) > 0 \\ & \text{and } \delta^K((s, K, m), \sigma_E, \sigma_A)((s', K', m')) > 0 \} \end{aligned}$$

Consider the following increasing sequence $(\text{Rank}_i)_{i \geq 0}$: $\text{Rank}_0 = (2^{F^K} \cap \mathcal{K}^{\text{AS}}) \times M$ consists of trivially winning positions and

$$\begin{aligned} \text{Rank}_{i+1} = \text{Rank}_i \cup \{ & (s, K, m) \mid K \in \mathcal{K}^{\text{AS}} \\ & \text{and } \exists \sigma_E \in \text{Allow}(K) \text{ s.t. } \text{Post}_{\sigma_E, \text{Move}(m)}(s, K, m) \cap \text{Rank}_i \neq \emptyset \} \end{aligned}$$

Let us denote by Rank^* the limit of the sequence $(\text{Rank}_i)_{i \geq 0}$. We claim that $\text{Rank}^* = \{(s, K, m) \mid K \in \mathcal{K}^{\text{AS}}\}$ and that φ is an almost-surely winning strategy for Eve from those positions when Adam plays accordingly to ψ .

The inclusion $\text{Rank}^* \subseteq \{(s, K, m) \mid K \in \mathcal{K}^{\text{AS}}\}$ is forced by the definition of Rank^* . The fact that φ is an almost-surely winning strategy for Eve from positions in Rank^* when Adam plays accordingly to ψ is a simple consequence of how $(\text{Rank}_i)_{i \geq 0}$ is defined and of Borel-Cantelli Lemma. Indeed from any configuration in Rank_i , there is a non null probability to reach a final state in the next i moves while playing φ against ψ and moreover the play surely stay inside Rank^* while playing φ against ψ : hence for any $k \geq 0$, the probability, that a play starting from Rank^* , in which Eve follows φ and Adam follows ψ , visits at most k time a final configuration is null. Therefore the probability of going infinitely often through a final state is 1, meaning that φ is almost-surely winning in Rank^* .

In order to prove the other inclusion, we let $X = \{(s, K, m) \mid K \in \mathcal{K}^{\text{AS}}\} \setminus \text{Rank}^*$ and assume by contradiction that $X \neq \emptyset$. By definition, for any element $(s, K, m) \in X$ we have that $\forall \sigma_E \in \text{Allow}(K)$, $\text{Post}_{\sigma_E, \text{Move}(m)}(s, K, m) \subseteq X$. This means in particular that following φ from such a configuration, and if Adam plays accordingly to ψ , then Eve surely loses as the play surely stay in X and $X \cap 2^{F^K} \times M = \emptyset$. Now we claim that the same holds if one replaces φ by any almost-surely winning strategy for Eve, leading to a contradiction. Indeed consider an almost-surely winning strategy φ^{AS} for Eve. Then we have the following fact (whose proof is omitted as it is exactly the same as the one of Fact 5).

Fact 9 *Let λ be a partial play consisting only of configurations in X . Assume that, for some strategy ψ' of Adam, λ is a possible partial play accordingly to both φ^{AS} and ψ' (more formally, $\Pr_{(s, K, m)}^{\varphi, \psi'}(\text{cone}(\lambda)) > 0$ where (s, K, m) denotes the initial configuration of λ). Then for any action $\sigma_E \in \Sigma_E$, one has $\varphi^{\text{AS}}(\lambda)(\sigma_E) > 0$ if and only if $\sigma_E \in \text{Allow}(K')$ where K' denotes Eve's knowledge in the last configuration of λ .*

Now one is ready to conclude. Assume Adam plays accordingly to ψ and Eve plays accordingly to some almost-surely winning strategy φ^{AS} . Then it follows from Fact 9 and definition of X that a play starting in X stays forever in X , hence never visits $F^K \times M$ and contradicting the hypothesis that φ^{AS} is almost-surely winning. Therefore $X = \emptyset$, which concludes the proof of Proposition 5 \square

Now one concludes the proof of Theorem 3 exactly as for the proof of Theorem 1. The 2-ExpTime hardness lower bound follows from the fact that it already holds for reachability objective (Theorem 2).

6 Discussion

The main contribution of this paper is to prove that one can decide whether Eve has an almost-surely winning strategy in a concurrent game with imperfect information equipped with a reachability objective or a Büchi objective.

A natural question is whether this result holds for other objectives, in particular for co-Büchi objectives. In a recent work [3], Baier *et al.* established undecidability of the emptiness problem for probabilistic Büchi automata on infinite words. Such an automaton can be simulated by a $1\frac{1}{2}$ -player imperfect information game: the states of the game are the one of the automaton, they are all equivalent for the player, and therefore an observation based strategy is an infinite word. Hence a pure (*i.e.* non-randomised) strategy in such a game coincide with an input word for the automaton. From this fact, Baier *et al.* derived that it is undecidable whether, in a $1\frac{1}{2}$ -player co-Büchi game with imperfect information, Eve has an almost-surely winning *pure* strategy.

One can also consider the stochastic-free version of this problem (an arena is **deterministic** iff $\delta(q, \sigma_E, \sigma_A)(q') \in \{0, 1\}$ for all $q, q', \sigma_E, \sigma_A$) and investigate whether one can decide if Eve has an almost-surely winning strategy in a *deterministic* game equipped with a co-Büchi objective. We believe that the $1\frac{1}{2}$ -player setting can be reduced to this new one, hence allowing to transfer undecidability results [15]. An even weaker model to consider is the stochastic-free model in which Adam has perfect information about the play [8].

It may happen that Eve has no almost-surely winning strategy while having a family $(\varphi_\varepsilon)_{0 < \varepsilon < 1}$ of strategies such that φ_ε ensures to win with probability at least $1 - \varepsilon$. Such a family is called **limit-surely winning**. Deciding existence of such families is a very challenging problem: indeed, in many practical situations, it is satisfying enough if one can control the risk of failing. Even if those questions have been solved for perfect information games [1], as far as we know, there has not been yet any result obtained in the imperfect information setting.

Even if the algorithms provided in this paper are "optimal", they are rather naive (checking all strategies for Eve may cost a lot in practice). Hence, one should look for fixpoint-based algorithms as the one studied in [8]: it would be of great help for a symbolic implementation, and it could also be a useful step toward a solution of the problem of finding limit-surely winning strategies. Note

that there are already efficient techniques and tools for finding *sure* winning strategies in subclasses of concurrent games with imperfect information [6, 5].

Acknowledgements. The authors want to warmly thank Krishnendu Chatterjee and Laurent Doyen for pointing out an important mistake in a previous version of this work originally published in ICALP’09 [14]. Indeed, in this previous work we additionally assumed that the players did not observe the actions they played and we proposed a variation of the constructions for this richer setting. It turned out that the constructions were wrong. We refer to [9] for more insight on this point as well as for an answer to this more general problem.

References

1. L. de Alfaro and T.A. Henzinger. Concurrent omega-regular games. In *Proceedings of LICS’00*, pages 141–154, 2000.
2. L. de Alfaro, T.A. Henzinger, and O. Kupferman. Concurrent reachability games. *Theoretical Computer Science*, 386(3):188–217, 2007.
3. C. Baier, N. Bertrand, and M. Größer. On decision problems for probabilistic büchi automata. In *Proceedings of FoSSaCS 2008*, volume 4962 of *LNCS*, pages 287–301. Springer, 2008.
4. N. Bertrand, B. Genest, and H. Gimbert. Qualitative Determinacy and Decidability of Stochastic Games with Signals. In *Proceedings of LICS 2009*, pages 319–328. IEEE, 2009.
5. D. Berwanger, K. Chatterjee, M. De Wulf, L. Doyen, and T.A. Henzinger. Alpaga: A tool for solving parity games with imperfect information. In *Proceedings of TACAS 2009*, volume 5505 of *LNCS*, pages 58–61. Springer, 2009.
6. D. Berwanger, K. Chatterjee, L. Doyen, T.A. Henzinger, and S. Raje. Strategy construction for parity games with imperfect information. In *Proceedings of CONCUR 2008*, volume 5201 of *LNCS*, pages 325–339. Springer, 2008.
7. K. Chatterjee. *Stochastic ω -Regular Games*. PhD thesis, University of California, 2007.
8. K. Chatterjee, L. Doyen, T.A. Henzinger, and J.-F. Raskin. Algorithms for omega-regular games with imperfect information. *Logical Methods in Computer Science*, 3(3), 2007.
9. K. Chatterjee, L. Doyen. Partial-Observation Stochastic Games: How to Win when Belief Fails. Preprint, <http://arxiv.org/abs/1107.2141>, July 2011.
10. F. Horn. Private communication. February 2009.
11. K. Chatterjee and T.A. Henzinger. Semiperfect-information games. In *Proceedings of FST&TCS 2005*, volume 3821 of *LNCS*, pages 1–18. Springer, 2005.
12. E. Grädel, W. Thomas, and Th. Wilke, editors. *Automata, Logics, and Infinite Games: A Guide to Current Research*, volume 2500 of *LNCS*. Springer, 2002.
13. Y. Gurevich and L. Harrington. Trees, automata, and games. In *Proceedings of STOC 1982*, pages 60–65, 1982.
14. V. Gripon and O. Serre. Qualitative Concurrent Games with Imperfect Information In *Proceedings of ICALP 2009*, volume 5556 of *LNCS*, pages 200–211. Springer, 2009.
15. F. Horn. Private communication. February 2009.
16. A. Paz. *Introduction to probabilistic automata*. Academic Press New York, 1971.

17. P.J. Ramadge and W.M. Wonham. Supervisory Control of a Class of Discrete Event Processes. *SIAM Journal on Control and Optimization*, 25:206, 1987.
18. J.H. Reif. The complexity of two-player games of incomplete information. *Journal of Computer and System Sciences*, 29(2):274-301, 1984.